Chlorination Studies II. The Reaction of Aqueous Hypochlorous Acid with -Amino Acids and Dipeptides.
Biochim. et Biophys. Acta, 313, 170 (1973).
By. W.E. Pereira, Y. Hoyano, R. Summons, V.A. Bacon and A.M. Duffield.

Chlorination Studies IV. The Reaction of Aqueous Hypochlorous Acid with Pyrimidine and Purine Bases.
Biochem. Biophys. Res. Commun., 53, 1195 (1973).
By Y. Hoyano, V. Bacon, R.E. Summons, W.E. Pereira, B. Halpern and A.M. Duffield.

Part C. EXTENSION OF THE THEORY OF MASS SPECTROMETRY BY COMPUTER

OBJECTIVES:

Part C of the DENDRAL effort, termed Meta-DENDRAL, aims at providing theory formation help for chemists interested in the mass spectrometric behavior of new classes of compounds. Our goals are necessarily long-range because theory formation by computer is itself an exciting, unsolved problem in computer science. We have chosen to explore this problem in the context of mass spectrometry in order to make frontier computer research results available to working scientists.

The problem of finding judgmental rules for use in a computer program is common to many biomedical computing projects, such as medical diagnosis and therapy recommendation programs. <See, for example, Shortliffe, et.al.> In order to give these programs the knowledge that makes them perform at acceptable levels, a medical expert is often asked to summarize his own knowledge of the problem area in rules that the program can use. The Meta-DENDRAL theory formation program is a paradigm for the kind of assistance that computers can give to the medical experts in this role. Programs of this sort can, first of all, provide the expert with an interpreted summary of a large collection of "hard" empirical data. Second, the program can suggest to the expert plausible rules that appear to explain major features of the data. Thus, the expert is able to assimilate large collections of data in the rules given to the computer. We believe that the meta-DENDRAL work is a useful model on which fruitful work in other biomedical problems can be based.

The over-all strategy of this research is to model the theory formation activity of scientists. We start with a set of empirical data which are known molecular structures and their associated mass spectra. By exploring the possible mechanistic explanations of each mass spectrum, the program is able to find a set of mechanisms that appear to be characteristic for the class of molecules. These characteristic mechanisms constitute the general mass spectrometry rules for the class, or a first-level theory for the class. Further refinements of the rules give more sophisticated restatements of the theory.

We have designed the programs in such a way as to provide useful results from the intermediate steps. The progress section discusses several sets of results that have been obtained, even though the entire program has not yet been completed.

PROGRESS:

In the past ten months (since January, 1973) the theory formation programs have seen significant application and significant new extensions. In addition, the work has been described in publications for both chemists and computer scientists.

Applications of Existing Programs.

The INTSUM program, for interpreting and summarizing the mass spectra of many known compounds of one class, was described in the previous annual report as essentially finished. In this last period we have used this program to help understand the mass spectrometry of several

classes of compounds, including estrogens, equilenins and other estrogenic steroids, androstanes, alkyl pregnanes, vinyl quinazalones, amino acids and aromatic acids. An article written for mass spectroscopists and soon to appear in Tetrahedron (Smith, et.al, enclosed) describes this program and its usefulness in understanding the previously unreported mass spectrometry of the equilenins. The amino acid and aromatic acid results are useful for interpreting the mass spectra taken from those fractions of urine (see Part B).

The INTSUM program is available to anyone who requests it, as stated in the article soon to appear. Because of the complexity of the program, we recommend that mass spectroscopists use this program on a network computer after they have collected a number of mass spectra from a class of compounds whose fragmentation mechanisms they wish to investigate.

Recent Extensions to Meta-DENDRAL.

In this last period significant progress has been made on the theory formation programs that use the interpreted summary of the data provided by the INTSUM program. A simple rule formation program, described previously (MI7), finds the characteristic mass spectrometry mechanisms for a class of compounds, assuming that the compounds exhibit regular behavior as a class. Recent work has removed the restriction that the compounds must behave as a class - important classes can be found by the program within the set of given compounds. The procedure was described in a paper for the Third International Joint Conference on Artificial Intelligence, which is enclosed. At the same time that the rule formation program looks for characteristic mechanisms, the class separation procedure refines the class of molecules that appear to behave uniformly (i.e., appear to exhibit most of the characteristic mechanisms).

Another important extension of the theory formation program makes the rule descriptions more general and less specific to the class of compounds studied. The mechanisms in the rules are now described generally in terms of the kinds of bonds that break, and not in terms of the precise relations of the bonds to the skeletal structure common to the class. For example, a rule is now stated as "Any bond that is the second bond from a nitrogen atom is likely to break", rather than "In the skeleton R1-C2-N3-C4-R5 the bond between atoms 1 and 2 and the bond between atoms 4 and 5 are both likely to break".

These general descriptions will allow much more freedom in the kinds of interpretations that can be placed on the INTSUM results. It is possible, for example, to alter the set of predicates used to describe bonds without altering the program.

The program can be conceptualized as a search program through the space of possible combinations of predicates. Some predicates describe the type of bond (e.g., 'single'), others describe the atoms joined by the bond (e.g., 'nitrogen', 'secondary'), and others describe the bonds and atoms next away from the bond that breaks. Some a priori heuristics limit consideration of complex predicates to chemically meaningful combinations, for example, by forbidding consideration of a single atom as both carbon and nitrogen. Other heuristics guide the process of expansion by forbidding a new predicate to be added to a description if its addition reduces the explanatory power of the existing description. For example, if a high average intensity is associated with breaking the

X-X bond in X-X-N and further specification of either of the X's reduces the average intensity, then the description is not changed.

In addition to the work just mentioned, a generative model of rule formation has been pursued by Carl Farrell in his dissertation work directed by Professor Feigenbaum and Dr. Buchanan. He has written a program which accepts, as input, descriptions of specific molecules and all the primitive actions that might explain the mass spectra of those molecules. The output of the program is a set of general situation-action rules that describe classes of molecules that seem to be characteristically show evidence of significant actions.

PLANS

In the following period we plan to increase the performance capabilities of the theory formation program in several ways.

1. Sample Selection.

The program's current strategy is to find the rules exhibited by most or all of the molecules in the initial sample. If the molecules are diverse, the rules will be diverse. Thus, we plan to add a preprocessor that can select a "simple" set of molecules for the rule formation to work with. For example, unbranched (straight-chain) compounds should be expected to present fewer complications for initial theory formation than highly branched compounds. The effects of the complicating features can be studied after the simple rules have been found.

2. Rule Clarification.

After simple rules have been found, we want the program to clarify the conditions under which the rules hold. By studying more complicated molecules, the program can find the simple rules that no longer hold for these cases. For example, we want the program to discover that terminal alpha carbons (as in CH3-X-N) are special. Or, the program should discover the effects of double bonds by examining new cases even though the molecules in the original set contained no double bonds.

3. Experimentation.

Because the original set of molecules contains the simpler examples from which it is easier to find characteristic mechanisms, the program will need to clarify rules in the way suggested under (2). For a human scientist, this means describing new experiments to perform that will help place limits on the range of applicability of the rules. Looking at additional arbitrary molecules may be helpful, but not as helpful as looking at the specific molecules that will resolve specific questions about the preliminary rule set.

4. Integration of Results.

When the program has examined two or more classes of molecules, it should be able to integrate the results into a common set of mechanisms (if any are common). The set of predicates used by the integration program may not have to be wider than the set used by the rule formation program, but one would expect the rules themselves to be more general. For example, integrating aliphatic amine and ether results should combine the separate alpha-cleavage rules (one with nitrogen, one

with oxygen) into a more general rule (specifying 'N or O', or 'heteroatom').

PART C REFERENCES (Published or submitted during this year)

D.H. Smith, B.G. Buchanan, W.C. White, E.A. Feigenbaum, C. Djerassi and J. Lederberg, "Applications of Artificial Intelligence for Chemical Inference X. INTSUM. A Data Interpretation Program as Applied to the Collected Mass Spectra of Estrogenic Steroids". Tetrahedron. In press.

B.G. Buchanan and N.S. Sridharan, "Analysis of Behavior of Chemical Molecules: Rule Formation on Non-Homogeneous Classes of Objects". In proceedings of the Third International Joint Conference on Artificial Intelligence, Stanford University (August, 1973). (Also Stanford Artificial Intelligence Project Memo No. 215.)

Related Publications

D. Michie and B.G. Buchanan, "Current Status of the Heuristic DENDRAL Program for Applying Artificial Intelligence to the Interpretation of Mass Spectra". August, 1973.

E.H. Shortliffe, S.G. Axline, B.G. Buchanan, T.C. Merrigan and S.N. Cohen, "An Artificial Intelligence Program to Advise Physicians Regarding Antimicrobial Therapy". Computers & Biomedical Research. In Press.

HUMAN SUBJECTS

HUMAN SUBJECTS

As a part of this research project, GC/MS analysis techniques will be applied to human body fluids in collaboration with clinical investigators, and blood and urine specimens will be collected from human subjects. Collection of VOIDED URINE SPECIMENS presents no risk to the patient. Collection of blood samples will not be taken solely for the purpose of this research but rather would be collected as part of a diagnostic procedure deemed necessary for clinical diagnosis.

The undersigned agrees to accept responsibility for the scientific and technical conduct of the project and for provision of required progress reports if a grant is awarded as the result of this application.

12/19/73
Date

Carl Djerassi
Principal Investigator

APPENDIX A

FIGURES 1-3

To Josh
From Dennis

Here is the additional background material you requested on ~~the~~ multiplet resolution and associated topics

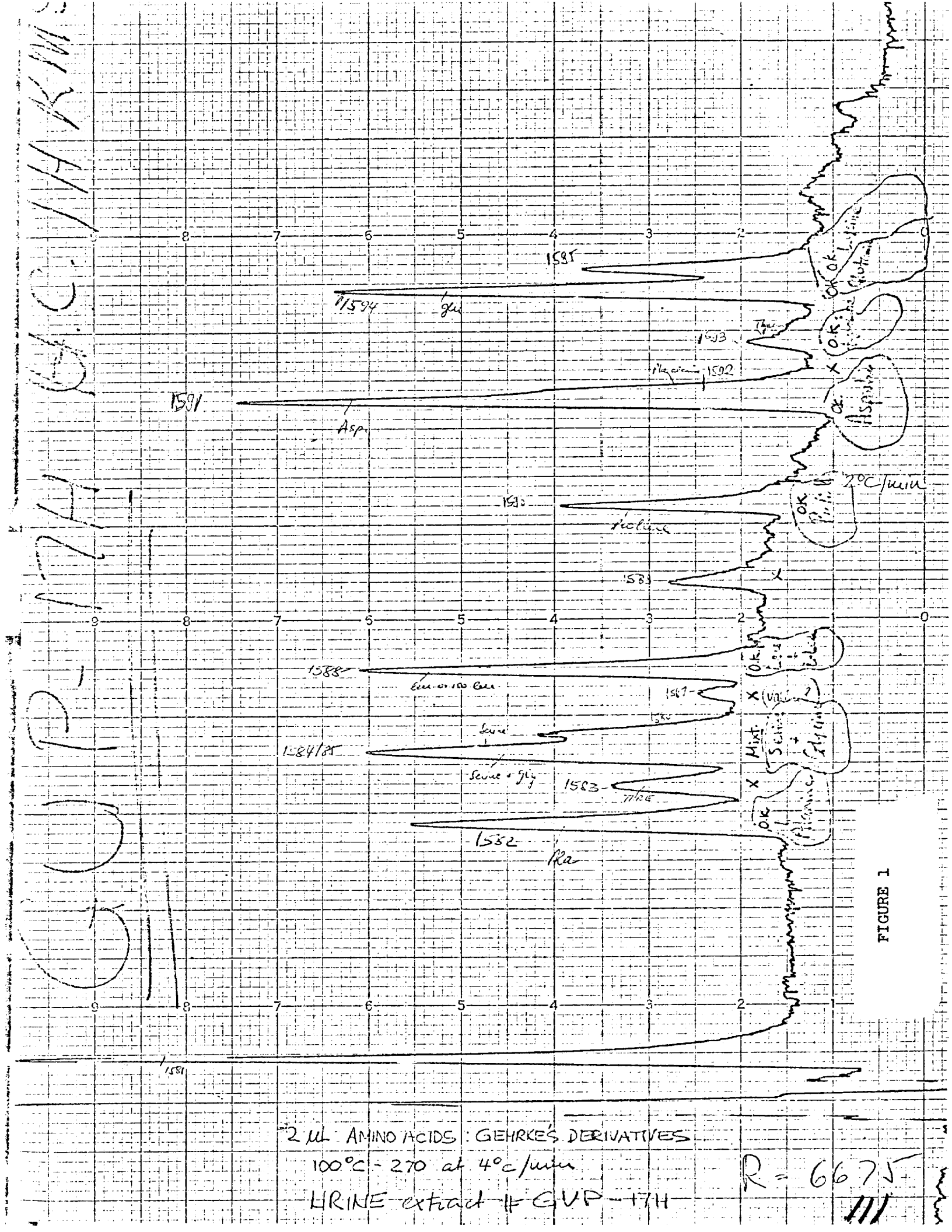
Dennis

do Statistics of Am

Table 1. Potential doublets in composition range C,H, + ... with mass differences (e) <50 mmu (0.05 amu).

C	H	N	O	E mmu	
0	0	0	0	0.0	
0	[illegible]	1	-1	23.4	..
0	[illegible]	2	-2	46.8	
0	-8	4	-3	-33.4	
1	-2	-3	2	-34.6	
1	0	-2	1	-11.2	..
1	2	-1	0	11.2	..
1	[illegible]	0	-1	35.6	
1	-6	2	-2	-44.7	
1	-6	3	-3	-21.3	
1	-4	4	-4	2.1	--
2	0	-4	2	-22.5	..
2	2	-3	1	0.9	--
2	4	-2	0	24.4	..
2	6	-1	-1	47.8	
2	-6	1	-2	-32.[illegible]	
2	-4	2	-3	-9.[illegible]	--
2	-2	3	-4	14.3	..
3	[illegible]	-4	1	13.1	..
3	6	-3	0	36.5	
3	-6	-1	-1	-43.7	
3	-4	0	-2	-20.3	..
3	-2	1	-3	3.1	--
[illegible]	[illegible]	2	-4	26.5	..
[illegible]	8	-4	0	48.7	
4	-4	-2	-1	-31.6	
4	-2	-1	-2	-9.2	--
4	0	0	-3	15.3	..
4	2	1	-4	38.7	

The unfinished tables will consider various ranges of the elements, ... consider isotopic nuclides, and will be sorted by e as well as ...



2 μL AMINO ACIDS : GEHRKE'S DERIVATIVES

100°C - 270 at 4°C/min

URINE extract + GVP - TH

R = 667

FIGURE 2

```
FILE DPH01594   USING REF FILE MPR01564
SOURCE IS urine-fraction:aminoacids(Gehrke technique
CMP ID IS GVP-17H
UNDEFINED     MASS OF MCL ION  DATE RUN 730313
RUN ON MAT 711 EXP DOWN SCAN
THRESHOLD= 4, PEAKS REJECTED   107 FOR AREA,     0 FOR WIDTH
SAMP RATE=10800 , MIN WIDTH= 2, MIN AREA= 18
(15.6 SECS, 1.49 DECS), TDEC=10.5
NUMBER OF PEAKS FOUND=170   ARRAY SPACE USED( 1750/ 8000)= 21.9%
 38 CALIB MASSES WITH LAST ONE=455.0,  6 MASSES ABOVE NOT FOUND
MISSED CALIBRATION MASSES:
281.0
******************  TABLE OF PROJECTION ERRORS **********************

 28.0      9.5   31.0     43.6   44.0      2.8   51.0     24.9   69.0    -0.3
 81.0      0.3   93.0      0.3  100.0      0.2  112.0     -0.5  119.0    -0.2
131.0     -0.2  143.0      0.1  155.0      0.1  162.0      0.1  169.0    -0.9
181.0      2.4  193.0     -2.3  205.0      1.5  212.0     -0.2  219.0    -0.9
231.0      4.1  243.0     -4.4  255.0      1.8  269.0     -0.9  293.0     1.5
305.0      0.4  319.0     -2.0  331.0      2.6  343.0     -2.2  355.0     2.6
367.0      0.6  381.0      1.1  393.0      5.1  405.0      4.4  417.0    -1.7
431.0      2.0  443.0     -4.0  455.0      6.6
MASS FILE MPH01594 HAS BEEN CREATED WITH 84 MASSES FOUND
SAMPLE BASE PEAK IS   3.7 VOLTS AND REF BASE PEAK IS 174.5 VOLTS
MATCHING TOLERANCE= 4.000 MMU FROM MASS  40 TO 700
MATCHING C 15 H 20 O  5 N  4 FL   6
   MASS        AREA     MMU ERROR     COMPOSITION
 41.03931      58.0       0.183       C  3 H  5
 43.05388       5.4      -0.899       C  3 H  7
 55.01820       3.5       1.153       C  1 H  1 N  3
                         -0.189       C  3 H  3 O  1
 55.05510      12.0       0.321       C  4 H  7
 56.06277      33.9       0.180       C  4 H  8
 57.07068      72.8       0.248       C  4 H  9
 64.01433       4.0      -1.721       C  1 H  4 O  3
                          1.863       C  2 H  2 FL   2
 68.96245       3.7
 68.97505       4.7
 69.03101       5.4      -1.698       C  2 H  3 N  3
                         -3.041       C  4 H  5 O  1
 70.03085       4.4       2.907       C  1 H  2 N  4
                          1.564       C  3 H  4 O  1 N  1
                         -3.405       H  4 N  2 FL   2
 75.00519       6.2      -1.682       H  1 O  2 N  3
                         -3.025       C  2 H  3 O  3
                          0.560       C  3 H  1 FL   2
                         -0.585       H  2 O  1 FL   3
 83.01006       5.8      -1.911       C  2 H  1 O  1 N  3
                         -3.253       C  4 H  3 O  2
                         -0.814       C  2 H  2 FL   3
 84.04382       6.0       0.230       C  2 H  4 N  4
                         -1.113       C  4 H  6 O  1 N  1
 85.02852      35.3       0.914       C  2 H  3 O  1 N  3
                         -0.429       C  4 H  5 O  2
```

		2.011	C 2 H 4 FL 3
96.00504	9.5	-0.833	C 1 H 4 C 5
		-2.170	C 2 O 1 N 4
		-3.513	C 4 H 2 O 2 N 1
		3.896	C 5 H 1 O 1 FL 1
		2.751	C 2 H 2 O 2 FL 2
		-1.073	C 2 H 1 N 1 FL 3
134.96208	3.3		
137.00035	4.6	1.639	C 5 H 1 O 3 N 2
		-2.383	C 10 H 1 O 1
		0.494	C 2 H 2 O 4 N 2 FL 1
		-2.185	C 5 O 1 N 3 FL 1
		-3.528	C 7 H 2 O 2 FL 1
		-3.330	C 2 H 1 O 2 N 3 FL 2
		2.934	H 1 O 2 N 2 FL 4
		-1.088	C 5 H 1 FL 4
		-0.890	N 3 FL 5
		-2.233	C 2 H 2 O 1 FL 5
139.02548	5.6	2.567	C 1 H 5 O 5 N 3
		-0.113	C 4 H 3 O 2 N 4
		-1.455	C 6 H 5 O 3 N 1
		-1.258	C 1 H 4 O 3 N 4 FL 1
		-2.600	C 3 H 6 O 4 N 1 FL 1
		2.129	C 7 H 3 N 1 FL 2
		2.327	C 2 H 2 N 4 FL 3
		0.984	C 4 H 4 O 1 N 1 FL 3
		-0.160	C 1 H 5 O 2 N 1 FL 4
		-3.985	C 1 H 4 N 2 FL 5
140.03549	3.2	2.065	C 4 H 4 O 2 N 4
		0.722	C 6 H 6 O 3 N 1
		-1.957	C 9 H 4 N 2
		0.920	C 1 H 5 O 3 N 4 FL 1
		-0.423	C 3 H 7 O 4 N 1 FL 1
		-3.102	C 6 H 5 O 1 N 2 FL 1
		3.162	C 4 H 5 O 1 N 1 FL 3
		-1.807	C 1 H 5 N 2 FL 5
152.03288	66.1	2.141	C 2 H 6 O 5 N 3
		-0.538	C 5 H 4 O 2 N 4
		-1.881	C 7 H 6 O 3 N 1
		-1.683	C 2 H 5 O 3 N 4 FL 1
		-3.025	C 4 H 7 O 4 N 1 FL 1
		1.704	C 8 H 4 N 1 FL 2
		1.902	C 3 H 3 N 4 FL 3
		0.559	C 5 H 5 O 1 N 1 FL 3
		0.757	H 4 O 1 N 4 FL 4
		-0.586	C 2 H 6 O 2 N 1 FL 4
153.03532	4.9	2.620	C 9 H 3 N 3
		1.278	C 11 H 5 O 1
		-3.254	C 2 H 7 O 5 N 3
		1.475	C 6 H 4 O 1 N 3 FL 1
		0.133	C 8 H 6 O 2 FL 1
		0.331	C 3 H 5 O 2 N 3 FL 2
		-3.691	C 8 H 5 N 1 FL 2
		-1.012	C 5 H 7 O 3 FL 2

		-3.494	C 3 H 4 N 4 FL 3
		2.573	C 6 H 5 FL 4
164.00092	3.5	3.877	C 5 O 3 N 4
		2.534	C 7 H 2 O 4 N 1
		-0.145	C 10 O 1 N 2
		2.732	C 2 H 1 O 4 N 4 FL 1
		1.389	C 4 H 3 O 5 N 1 FL 1
		-1.290	C 7 H 1 O 2 N 2 FL 1
		-2.435	C 4 H 2 O 3 N 2 FL 2
		-3.580	C 1 H 3 O 4 N 2 FL 3
		3.829	C 2 H 2 O 3 N 1 FL 4
		1.150	C 5 N 2 FL 4
173.98967	6.3	1.963	C 5 O 3 N 2 FL 2
		-2.059	C 10 O 1 FL 2
		0.818	C 2 H 1 O 4 N 2 FL 3
		-3.204	C 7 H 1 O 2 FL 3
		-3.006	C 2 O 2 N 3 FL 4
		3.258	C 2 N 2 FL 6
		-0.764	C 5 FL 6
174.99890	11.4	-2.896	C 7 H 1 O 3 N 3
		0.491	C 13 FL 1
		3.368	C 5 H 1 O 3 N 2 FL 2
		-0.654	C 10 H 1 O 1 FL 2
		2.223	C 2 H 2 O 4 N 2 FL 3
		-0.456	C 5 O 1 N 3 FL 3
		-1.799	C 7 H 2 O 2 FL 3
		-1.601	C 2 H 1 O 2 N 3 FL 4
		-2.943	C 4 H 3 O 3 FL 4
		0.641	C 5 H 1 FL 6
178.98990	3.5	0.599	C 7 O 3 N 2 FL 1
		-3.423	C 12 O 1 FL 1
		-0.546	C 4 H 1 O 4 N 2 FL 2
		-1.691	C 1 H 2 O 5 N 2 FL 3
		1.894	C 2 O 2 N 2 FL 5
		-2.128	C 7 FL 5
		-3.273	C 4 H 1 O 1 FL 6
180.02676	87.2	-1.582	C 6 H 4 O 3 N 4
		-2.925	C 8 H 6 O 4 N 1
		1.805	C 12 H 3 N 1 FL 1
		-2.727	C 3 H 5 O 4 N 4 FL 1
		3.339	C 6 H 6 O 4 FL 2
		2.003	C 7 H 2 N 4 FL 2
		0.660	C 9 H 4 O 1 N 1 FL 2
		-3.872	H 6 O 5 N 4 FL 2
		3.537	C 1 H 5 O 4 N 3 FL 3
		2.194	C 3 H 7 O 5 FL 3
181.03192	12.5	2.973	C 12 H 5 O 2
		3.171	C 7 H 4 O 2 N 3 FL 1
		-0.851	C 12 H 4 N 1 FL 1
		1.828	C 9 H 6 O 3 FL 1
		2.026	C 4 H 5 O 3 N 3 FL 2
		0.683	C 6 H 7 O 4 FL 2
		-0.653	C 7 H 3 N 4 FL 2
		-1.996	C 9 H 5 O 1 N 1 FL 2

		0.881	C 1 H 6 O 4 N 3 FL 3
		-0.462	C 3 H 8 O 5 FL 3
190.99983	7.6	3.119	C 7 H 1 O 4 N 3
		1.776	C 9 H 3 O 5
		-0.903	C 12 H 1 O 2 N 1
		1.974	C 4 H 2 O 5 N 3 FL 1
		-0.705	C 7 O 2 N 4 FL 1
		-2.048	C 9 H 2 O 3 N 1 FL 1
		-1.850	C 4 H 1 O 3 N 4 FL 2
		-3.193	C 6 H 3 O 4 N 1 FL 2
		1.537	C 10 N 1 FL 3
		-2.995	C 1 H 2 O 4 N 4 FL 3
196.95901	3.2		
198.03705	100.0	2.670	C 15 H 4 N 1
		-1.862	C 6 H 6 O 4 N 4
		-3.204	C 8 H 8 O 5 N 1
		2.868	C 10 H 3 N 4 FL 1
		1.525	C 12 H 5 O 1 N 1 FL 1
		-3.006	C 3 H 7 O 5 N 4 FL 1
		3.060	C 6 H 8 O 5 FL 2
		1.723	C 7 H 4 O 1 N 4 FL 2
		0.380	C 9 H 6 O 2 N 1 FL 2
		3.258	C 1 H 7 O 5 N 3 FL 3
199.03798	3.7	-0.190	C 10 H 5 O 2 N 3
		2.489	C 7 H 7 O 5 N 2
		-1.533	C 12 H 7 O 3
		-1.335	C 7 H 6 O 3 N 3 FL 1
		-2.678	C 9 H 8 O 4 FL 1
		2.052	C 13 H 5 FL 2
		-2.480	C 4 H 7 O 4 N 3 FL 2
		-3.823	C 6 H 9 O 5 FL 2
		2.250	C 8 H 4 N 3 FL 3
		0.907	C 10 H 6 O 1 FL 3
199.98764	7.7	1.838	C 9 O 4 N 2
		-2.184	C 14 O 2
		0.693	C 6 H 1 O 5 N 2 FL 1
		-3.329	C 11 H 1 O 3 FL 1
		-3.131	C 6 O 3 N 3 FL 2
		3.133	C 4 O 3 N 2 FL 4
		-0.889	C 9 O 1 FL 4
		1.988	C 1 H 1 O 4 N 2 FL 5
		-2.034	C 6 H 1 O 2 FL 5
		-1.836	C 1 O 2 N 3 FL 6
200.99631	26.7	2.681	C 9 H 1 O 4 N 2
		-1.341	C 14 H 1 O 2
		-1.143	C 9 O 2 N 3 FL 1
		1.536	C 6 H 2 O 5 N 2 FL 1
		-2.486	C 11 H 2 O 3 FL 1
		-2.288	C 6 H 1 O 3 N 3 FL 2
		-3.631	C 8 H 3 O 4 FL 2
		1.099	C 12 FL 3
		-3.433	C 3 H 2 O 4 N 3 FL 3
		3.976	C 4 H 1 O 3 N 2 FL 4
217.99628	4.4	-2.778	C 12 O 2 N 3

		-0.099	C 9 H 2 C 5 N 2
		-3.923	C 9 H 1 C 3 N 3 FL 1
		3.486	C 10 C 2 N 2 FL 2
		-0.536	C 15 FL 2
		2.341	C 7 H 1 C 3 N 2 FL 3
		-1.681	C 12 H 1 C 1 FL 3
		1.196	C 4 H 2 C 4 N 2 FL 4
		-1.483	C 7 C 1 N 3 FL 4
		-2.826	C 9 H 2 C 2 FL 4
219.99133	7.9	0.443	C 12 C 3 N 2
		-0.702	C 9 H 1 C 4 N 2 FL 1
		-1.846	C 6 H 2 C 5 N 2 FL 2
		1.738	C 7 C 2 N 2 FL 4
		-2.284	C 12 FL 4
		0.593	C 4 H 1 C 3 N 2 FL 5
		-3.429	C 9 H 1 C 1 FL 5
		-0.551	C 1 H 2 C 4 N 2 FL 6
		-3.231	C 4 C 1 N 3 FL 6
223.98842	10.9	2.617	C 11 C 4 N 2
		1.472	C 8 H 1 C 5 N 2 FL 1
		-2.550	C 13 H 1 C 3 FL 1
		-2.352	C 8 C 3 N 3 FL 2
		-3.695	C 10 H 2 C 4 FL 2
		-3.497	C 5 H 1 C 4 N 3 FL 3
		3.912	C 6 C 3 N 2 FL 4
		-0.110	C 11 C 1 FL 4
		2.767	C 3 H 1 C 4 N 2 FL 5
		-1.255	C 8 H 1 C 2 FL 5
225.02393	6.2	3.802	C 14 H 1 N 4
		-2.072	C 7 H 5 C 5 N 4
		3.994	C 10 H 6 C 5 FL 1
		2.657	C 11 H 2 C 1 N 4 FL 1
		1.315	C 13 H 4 C 2 N 1 FL 1
		1.512	C 8 H 3 C 2 N 4 FL 2
		0.170	C 10 H 5 C 3 N 1 FL 2
		-2.510	C 13 H 3 N 2 FL 2
		0.368	C 5 H 4 C 3 N 4 FL 3
		-0.975	C 7 H 6 C 4 N 1 FL 3
226.03322	10.2	-0.603	C 7 H 6 C 5 N 4
		2.784	C 13 H 5 C 2 N 1 FL 1
		2.982	C 8 H 4 C 2 N 4 FL 2
		1.639	C 10 H 6 C 3 N 1 FL 2
		-1.040	C 13 H 4 N 2 FL 2
		1.837	C 5 H 5 C 3 N 4 FL 3
		0.494	C 7 H 7 C 4 N 1 FL 3
		-2.185	C 10 H 5 C 1 N 2 FL 3
		0.692	C 2 H 6 C 4 N 4 FL 4
		-0.651	C 4 H 8 C 5 N 1 FL 4
236.99301	3.3	-0.625	C 12 H 1 C 4 N 2
		-1.770	C 9 H 2 C 5 N 2 FL 1
		1.815	C 10 C 2 N 2 FL 3
		-2.207	C 15 FL 3
		0.670	C 7 H 1 C 3 N 2 FL 4
		-3.352	C 12 H 1 C 1 FL 4

		-0.475	C 4 H 2 O 4 N 2 FL 5
		-3.154	C 7 O 1 N 3 FL 5
		-1.620	C 1 H 3 O 5 N 2 FL 6
243.99310	5.8	2.209	C 14 O 3 N 2
		1.064	C 11 H 1 O 4 N 2 FL 1
		-2.760	C 11 O 2 N 3 FL 2
		-0.081	C 8 H 2 O 5 N 2 FL 2
		-3.905	C 8 H 1 O 3 N 3 FL 3
		3.504	C 9 O 2 N 2 FL 4
		-0.518	C 14 FL 4
		2.359	C 6 H 1 O 3 N 2 FL 5
		-1.663	C 11 H 1 O 1 FL 5
		1.214	C 3 H 2 O 4 N 2 FL 6
245.00569	5.3	-1.586	C 10 H 3 O 5 N 3
		3.144	C 14 O 1 N 3 FL 1
		1.999	C 11 H 1 O 2 N 3 FL 2
		0.656	C 13 H 3 O 3 FL 2
		0.854	C 8 H 2 O 3 N 3 FL 3
		-0.489	C 10 H 4 O 4 FL 3
		-1.825	C 11 N 4 FL 3
		-3.168	C 13 H 2 O 1 N 1 FL 3
		-0.291	C 5 H 3 O 4 N 3 FL 4
		-1.634	C 7 H 5 O 5 FL 4
247.98857	4.2	2.766	C 13 O 4 N 2
		1.621	C 10 H 1 O 5 N 2 FL 1
		-2.401	C 15 H 1 O 3 FL 1
		-2.203	C 10 O 3 N 3 FL 2
		-3.546	C 12 H 2 O 4 FL 2
		-3.348	C 7 H 1 O 4 N 3 FL 3
		0.039	C 13 O 1 FL 4
		2.916	C 5 H 1 O 4 N 2 FL 5
		-1.106	C 10 H 1 O 2 FL 5
		-0.908	C 5 O 2 N 3 FL 6
248.99155	4.7	-2.087	C 13 H 1 O 4 N 2
		-3.232	C 10 H 2 O 5 N 2 FL 1
		3.032	C 8 H 2 O 5 N 1 FL 3
		0.353	C 11 O 2 N 2 FL 3
		3.230	C 3 H 1 O 5 N 4 FL 4
		-0.792	C 8 H 1 O 3 N 2 FL 4
		-1.937	C 5 H 2 O 4 N 2 FL 5
		-3.082	C 2 H 3 O 5 N 2 FL 6
254.10146	39.5	-0.052	C 10 H 14 O 4 N 4
		-1.395	C 12 H 16 O 5 N 1
		-1.197	C 7 H 15 O 5 N 4 FL 1
		3.533	C 11 H 12 O 1 N 4 FL 2
		2.190	C 13 H 14 O 2 N 1 FL 2
		2.388	C 8 H 13 O 2 N 4 FL 3
		1.045	C 10 H 15 O 3 N 1 FL 3
		-1.634	C 13 H 13 N 2 FL 3
		1.243	C 5 H 14 O 3 N 4 FL 4
		-0.100	C 7 H 16 O 4 N 1 FL 4
255.10947	3.3	0.143	C 10 H 15 O 4 N 4
		-1.200	C 12 H 17 O 5 N 1
		-3.879	C 15 H 15 O 2 N 2

M - CO_2

		-1.002	C 7 H 16 O 5 N 4 FL 1
		3.728	C 11 H 13 O 1 N 4 FL 2
		2.385	C 13 H 15 O 2 N 1 FL 2
		2.583	C 8 H 14 O 2 N 4 FL 3
		1.240	C 10 H 16 O 3 N 1 FL 3
		-1.439	C 13 H 14 N 2 FL 3
		1.438	C 5 H 15 O 3 N 4 FL 4
255.99071	4.4	3.833	C 10 O 5 N 4
		-0.189	C 15 O 3 N 2
		-1.334	C 12 H 1 O 4 N 2 FL 1
		-2.479	C 9 H 2 O 5 N 2 FL 2
		3.785	C 7 H 2 O 5 N 1 FL 4
		1.106	C 10 O 2 N 2 FL 4
		-2.916	C 15 FL 4
		3.983	C 2 H 1 O 5 N 4 FL 5
		-0.039	C 7 H 1 O 3 N 2 FL 5
		-1.184	C 4 H 2 O 4 N 2 FL 6
256.94409	19.5		
261.98169	7.4	-0.900	C 11 O 4 N 2 FL 2
		-2.045	C 8 H 1 O 5 N 2 FL 3
		0.395	C 6 O 3 N 2 FL 6
		-3.627	C 11 O 1 FL 6
262.99121	19.3	2.023	C 14 O 3 N 2 FL 1
		0.878	C 11 H 1 O 4 N 2 FL 2
		-2.946	C 11 O 2 N 3 FL 3
		-0.267	C 8 H 2 O 5 N 2 FL 3
		3.318	C 9 O 2 N 2 FL 5
		-0.704	C 14 FL 5
		2.173	C 6 H 1 O 3 N 2 FL 6
		-1.849	C 11 H 1 O 1 FL 6
269.98486	4.2	1.101	C 13 H 1 O 5 N 1 FL 1
		1.299	C 8 O 5 N 4 FL 2
		-2.723	C 13 O 3 N 2 FL 2
		-3.868	C 10 H 1 O 4 N 2 FL 3
		3.541	C 11 O 3 N 1 FL 4
		2.396	C 8 H 1 O 4 N 1 FL 5
		2.594	C 3 O 4 N 4 FL 6
		1.251	C 5 H 2 O 5 N 1 FL 6
		-1.428	C 8 O 2 N 2 FL 6
273.98608	10.7	-2.749	C 14 O 4 N 3
		-3.894	C 11 H 1 O 5 N 3 FL 1
		3.515	C 12 O 4 N 2 FL 2
		2.370	C 9 H 1 O 5 N 2 FL 3
		-1.652	C 14 H 1 O 3 FL 3
		-1.454	C 9 O 3 N 3 FL 4
		-2.797	C 11 H 2 O 4 FL 4
		-2.599	C 6 H 1 O 4 N 3 FL 5
		-3.942	C 8 H 3 O 5 FL 5
		0.788	C 12 O 1 FL 6
********** MISSED CALIBRATION MASS=281.0 **********			
281.08472	4.7	-3.910	C 11 H 13 O 5 N 4
		3.499	C 12 H 12 O 4 N 3 FL 1
		2.157	C 14 H 14 O 5 FL 1
		0.820	C 15 H 10 O 1 N 4 FL 1

		2.354	C 9 H 13 O 5 N 3 FL 2
		-0.325	C 12 H 11 O 2 N 4 FL 2
		-1.668	C 14 H 13 O 3 N 1 FL 2
		-1.470	C 9 H 12 O 3 N 4 FL 3
		-2.812	C 11 H 14 O 4 N 1 FL 3
		1.917	C 15 H 11 N 1 FL 4
282.09814	11.1	1.728	C 11 H 14 O 5 N 4
		-3.439	C 13 H 15 O 4 N 2 FL 1
		3.970	C 14 H 14 O 3 N 1 FL 2
		2.825	C 11 H 15 O 4 N 1 FL 3 M-OB
		0.146	C 14 H 13 O 1 N 2 FL 3
		3.023	C 6 H 14 O 4 N 4 FL 4
		1.681	C 8 H 16 O 5 N 1 FL 4
		-0.999	C 11 H 14 O 2 N 2 FL 4
		-2.144	C 8 H 15 O 3 N 2 FL 5
282.99707	5.3	1.422	C 14 H 1 O 3 N 2 FL 2
		0.277	C 11 H 2 O 4 N 2 FL 3
		-2.403	C 14 O 1 N 3 FL 3
		-3.547	C 11 H 1 O 2 N 3 FL 4
		-0.868	C 8 H 3 O 5 N 2 FL 4
		3.861	C 12 O 1 N 2 FL 5
		2.717	C 9 H 1 O 2 N 2 FL 6
		-1.305	C 14 H 1 FL 6
285.98682	9.3	-2.073	C 15 O 4 N 3
		-3.218	C 12 H 1 O 5 N 3 FL 1
		3.046	C 10 H 1 O 5 N 2 FL 3
		-0.976	C 15 H 1 O 3 FL 3
		-0.778	C 10 O 3 N 3 FL 4
		-2.121	C 12 H 2 O 4 FL 4
		-1.923	C 7 H 1 O 4 N 3 FL 5
		-3.266	C 9 H 3 O 5 FL 5
		1.464	C 13 O 1 FL 6
		-3.068	C 4 H 2 O 5 N 3 FL 6
290.98560	4.6	1.376	C 15 O 4 N 2 FL 1
		0.231	C 12 H 1 O 5 N 2 FL 2
		-3.593	C 12 O 3 N 3 FL 3
		2.671	C 10 O 3 N 2 FL 5
		-1.351	C 15 O 1 FL 5
		1.526	C 7 H 1 O 4 N 2 FL 6
		-2.496	C 12 H 1 O 2 FL 6
293.98828	4.7	0.548	C 15 O 3 N 2 FL 2
		-0.597	C 12 H 1 O 4 N 2 FL 3
		-1.742	C 9 H 2 O 5 N 2 FL 4
		1.843	C 10 O 2 N 2 FL 6
		-2.179	C 15 FL 6
297.97510	3.2	3.794	C 15 O 5 FL 2
		-1.175	C 12 O 4 N 1 FL 4
		-2.320	C 9 H 1 O 5 N 1 FL 5
		-2.122	C 4 O 5 N 4 FL 6
311.97388	4.9	-0.576	C 15 O 5 N 1 FL 2
		0.719	C 10 O 4 N 1 FL 6
312.98804	11.1	1.882	C 15 O 3 N 2 FL 3
		0.737	C 12 H 1 O 4 N 2 FL 4
		-3.087	C 12 O 2 N 3 FL 5

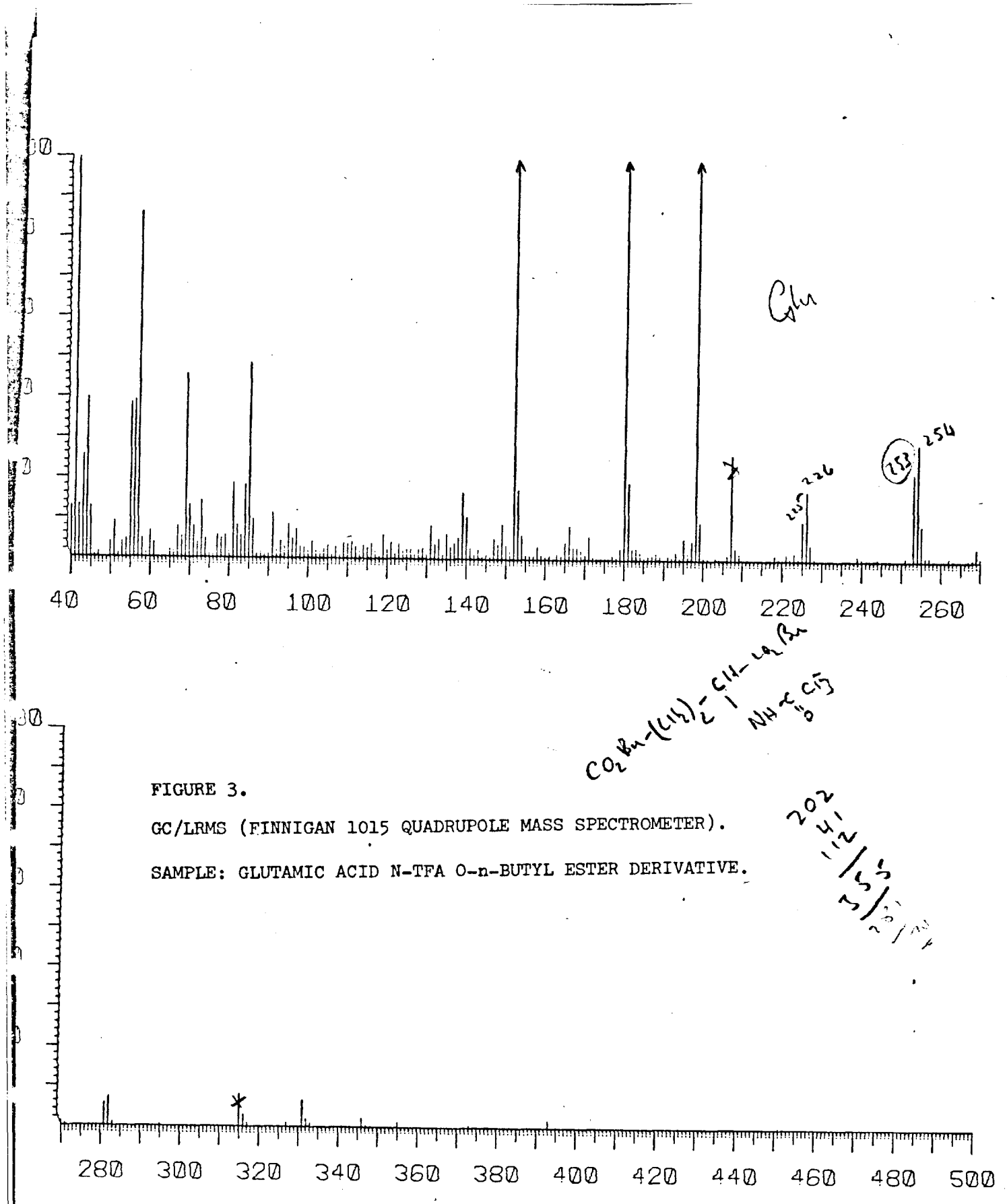


FIGURE 3.

GC/LRMS (FINNIGAN 1015 QUADRUPOLE MASS SPECTROMETER).

SAMPLE: GLUTAMIC ACID N-TFA O-n-BUTYL ESTER DERIVATIVE.

APPENDIX B

LETTERS OF INTEREST

STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

DEPARTMENT OF CHEMISTRY

December 17, 1973

Professor Carl Djerassi
Department of Chemistry
Stanford University
Stanford, California 94305

Dear Carl:

I am writing to indicate the anticipated use of mass spectral facilities by my research group in the foreseeable future. As has been true in the past, we plan to utilize both GC/HRMS and simple HRMS for various purposes, especially 1) the determination of structure of enzymic cyclization products, including members of the lanosterol class, derived from squalene oxide-like substrates, the purpose being the elucidation of the mechanism of enzymic steroid synthesis, and 2) the characterization and confirmation of structures of intermediates in the synthesis of natural products, including polycyclic terpenoids, alkaloids of physiological interest, and nucleosides,and 3) identification and/or structure determination of organic materials employed in our organic-inorganic program devoted to nitrogen fixation and related processes.

Very truly yours,

EEvT/jlb

121

DATE: December 3, 73

TO : Carl Djerassi

FROM : Keith Hodgson

SUBJECT: Response to inquiry about GC/HRMS facility

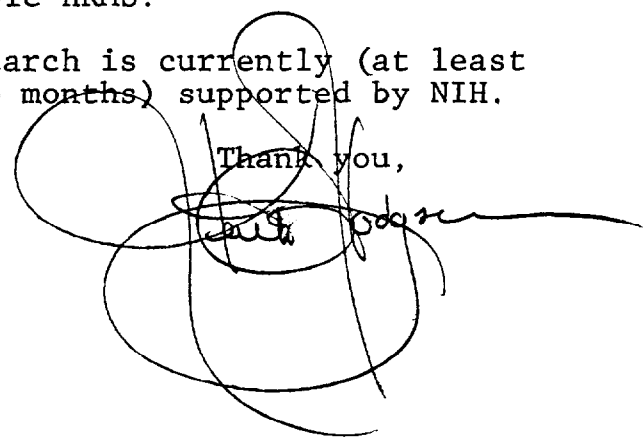
In response to your three questions concerning the potential use of upgraded GC/HRMS facilities:

1. Yes, especially in the study of certain biological ligands and lower molecular weight ligand-metal complexes.

2. Potential use of the facility might run in the range of 8-10 samples per year most of which probably would be handled most easily by simple HRMS.

3. No, no research is currently (at least for the next 6 months) supported by NIH.

Thank you,



STANFORD UNIVERSITY • OFFICE MEMORANDUM • STANFORD UNIVERSITY • OFFICE MEMORANDUM

DATE: December 3, 1973

TO : Professor Carl Djerassi

FROM : James P. Collman
Professor of Chemistry

SUBJECT:

Please excuse our belated response to your inquiry of November 20 concerning a potential upgrading of mass spectrometry facilities. The service you mention in your memo of the 20th would be valuable to us. We would have significant use for the GC/HRMS for a project dealing with models for cytochrome P_{450} based monooxygenases currently supported by the NIH.

JPC:lb

James P. Collman

DATE: December 13, 1973

To : Professor Djerassi

FROM : Professor Harry S. Mosher

SUBJECT: Your proposal to the NIH

On our NIH Grant on the investigation of animal toxins we have been studying natural products from the skin of Central American frogs (atelopidtoxin) and some products from marine animals (nudibranchs) as well as some new choline esters isolated from the hypobranchial gland of various sea snails. Some, if not all, of these are mixtures. Obviously the new capabilities of the mass spectrometry laboratory would be of value to me. I expect only occasional use of HRMS and GC/HRMS, but on these occasions these techniques would be very important.

Harry S. Mosher

• STANFORD UNIVERSITY • OFFICE MEMORANDUM • STANFORD UNIVERSITY • OFFICE MEMORANDUM • STANFORD UNIVERSITY • OFFICE MEMORANDUM

DATE: 14 December 1973

TO : C. Djerassi

FROM : W. S. Johnson

SUBJECT:

The contemplated new facility for high resolution mass spectrometry and combined gas chromatography/high resolution mass spectrometry would be of extreme value to our research program concerning the non-enzymic biogenetic-like cyclization of polyolefines, a project which is presently supported by NIH Grant AM 3787-14. If this facility were to become available, we would expect to use it extensively in the analysis of product mixtures of the aforementioned cyclizations. We estimate that our need for the gas chromatographic capability would be about 20% of the total need for the mass spectrometry service.

W. S. Johnson cs